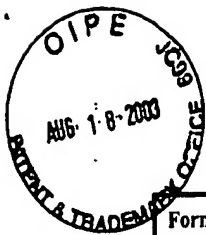
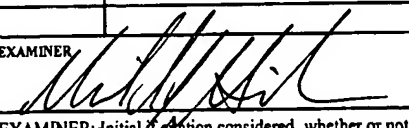




| | | | | | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------|---------|----------------------------------|-------------------------------|
| Form PTO-1449 U.S. DEPARTMENT OF COMMERCE (Modified) Patent and Trademark Office | | Attorney's Docket Number 16-358 | | Application Number 10/600,083 | |
| INFORMATION DISCLOSURE CITATION | | Applicant Surajit Chaudhuri, et al. | | Filed June 20, 2003 | |
| | | Examiner | | Art Unit | |
| Sheet 1 of 2 | | | | | |
| U.S. PATENT DOCUMENTS | | | | | |
| EXAMINER INITIALS | DOCUMENT NUMBER | DATE | NAME | CLASS | FILING DATE IF APPROPRIATE |
| | | | | | |
| FOREIGN PATENT DOCUMENTS | | | | | |
| | DOCUMENT NUMBER | DATE | COUNTRY | CLASS | SUBCLASS |
| | | | | | |
| OTHER DOCUMENTS | | | | | |
| MCH | R. Ananthakrishna, S. Chaudhuri and V. Ganti. <u>Eliminating Fuzzy Duplicates in Data Warehouses</u> . In <i>Proceedings of VLDB</i> , Hong Kong, 2002. | | | | |
| | A.N. Arslan, O. Egecioglu and P.A. Pevzner. <u>A New Approach to Sequence Comparison: Normalized Local Alignment</u> . <i>Bioinformatics</i> , 17(4):327-337, 2001 | | | | |
| | J.A. Aslam, K. Pelekhev and D. Rus. <u>Static and Dynamic Information Organization with Star Clusters</u> . <i>CIKM</i> 1998, pages 208-217. | | | | |
| | J.A. Aslam, K. Pelekhev and D. Rus. <u>A Practical Clustering Algorithm for Static and Dynamic Information Organization</u> . <i>ACM-SIAM Symposium on Discrete Algorithms</i> , 1999. | | | | |
| | V. Borkar, K. Deshmukh and S. Sarawagi. <u>Automatic Segmentation of Text Into Structured Records</u> . In <i>Proceedings of ACM SIGMOD Conference</i> , Santa Barbara, CA, May 2001. | | | | |
| | A. Broder, S. Glassman, M. Manasse and G. Zweig. <u>Syntactic Clustering of the Web</u> . In <i>Proc. Sixth Int'l. World Wide Web Conference, World Wide Web Consortium</i> , Cambridge, pages 391-404, 1997. | | | | |
| | K. Beyer, J. Goldstein, R. Ramakrishnan and U. Shaft. <u>When is "Nearest Neighbor" Meaningful?</u> <i>International Conference on Database Theory</i> , pages 217-235. January 1999. | | | | |
| | S. Chaudhuri, V. Ganjam, V. Ganti and R. Motwani. <u>Robust and Efficient Fuzzy Match for Online Data Cleaning</u> . In <i>Proceedings of ACM SIGMOD</i> , San Diego, CA June 2003. | | | | |
| | W. Cohen. <u>Integration of Heterogenous Databases Without Common Domains Using Queries Based on Textual Similarity</u> . In <i>Proceedings of ACM SIGMOD</i> , pages 201-212, Seattle, WA June 1998. | | | | |
| | R. Forino. <u>Data e.quality: A behind the Scenes Perspective on Data Cleansing</u> . http://www.dmreview.com/ , March 2001. | | | | |
| | H. Galhardas, D. Florescu, D. Shasha, E. Simon and C. Saita. <u>Declarative Data Cleaning: Language, Model, and Algorithms</u> . In <i>Proceedings of the 27th International Conference on Very Large Databases</i> , pages 371-380, Roma, Italy, September 11-14, 2001. | | | | |
| | H. Galhardas, D. Florescu, D. Shasha and E. Simon. <u>An Extensible Framework for Data Cleaning</u> . In <i>ACM SIGMOD</i> , May 1999. | | | | |
| | L. Gravano, P. Ipeirotis, H.V. Jagadish, N. Kousas, S. Muthukrishnan and D. Srivastava. <u>Approximate String Joins in a Database (Almost) For Free</u> . In <i>Proceedings of the VLDB</i> 2001. | | | | |
| | V. Ganti, J. Gehrke and R. Ramakrishnan. <u>Cactus-Clustering Categorical Data Using Summaries</u> . In <i>Proceedings of the ACM SIGKDD Fifth International Conference on Knowledge Discovery in Databases</i> , pages 73-83, August 15-18, 1999 | | | | |
| | D. Gibson, J. Kleinberg and P. Raghavan. <u>Clustering Categorical Data: An Approach Based on Dynamical Systems</u> . <i>VLDB</i> 1998, New York City, New York, August 24-27 | | | | |
| | S. Guha, R. Rastogi and K. Shim. <u>Rock: A Robust Clustering Algorithm for Categorical Attributes</u> . In <i>Proceedings of the IEEE International Conference on Data Engineering</i> , Sydney, March 1999. | | | | |
| | Y. Huhtala, J. Karkkainen, P. Porkka and H. Toivonen. <u>Efficient Discovery of Functional and Approximate Dependencies Using Partitions</u> . In <i>Proceedings of the 14th International Conference on Data Engineering (ICDE)</i> , pages 392-401, Orlando, Florida, February 1998. | | | | |
| EXAMINER | DATE CONSIDERED | | | | |
| | 12/2/05 | | | | |
| EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP § 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant. | | | | | |



| | | | | | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------|---------|----------------------------------|-------------------------------|
| Form PTO-1449 U.S. DEPARTMENT OF COMMERCE (Modified) Patent and Trademark Office | | Attorney's Docket Number 16-358 | | Application Number 10/600,083 | |
| INFORMATION DISCLOSURE CITATION | | Applicant Surajit Chaudhuri, et al. | | Filed June 20, 2003 | |
| | | Examiner | | Art Unit | |
| Sheet 2 of 2 | | | | | |
| U.S. PATENT DOCUMENTS | | | | | |
| EXAMINER INITIALS | DOCUMENT NUMBER | DATE | NAME | CLASS | FILING DATE IF APPROPRIATE |
| | | | | | |
| | | | | | |
| FOREIGN PATENT DOCUMENTS | | | | | |
| | DOCUMENT NUMBER | DATE | COUNTRY | CLASS | SUBCLASS |
| | | | | | |
| OTHER DOCUMENTS | | | | | |
| MATCH | M. Hernandez and S. Stolfo. <u>The Merge/Purge Problem for Large Databases</u> . In <i>Proceedings of the ACM SIGMOD</i> , pages 127-138, San Jose, CA May 1995 | | | | |
| | J. Madhavan, P. Bernstein, E. Rahm. <u>Generic Schema Matching with Cupid</u> . <i>VLDB 2001</i> , pages 49-58, Roma, Italy. | | | | |
| | A. Monge and C. Elkan. <u>The Field Matching Problem: Algorithms and Applications</u> . In <i>Proceedings of the Second International Conference on Knowledge Discovery and Databases (KDD)</i> , 1996. | | | | |
| | A. Monge and C. Elkan. <u>An Efficient Domain-Independent Algorithm for Detecting Approximately Duplicate Database Records</u> . In <i>Proceedings of the SIGMOD Workshop on Data Mining and Knowledge Discovery</i> , Tucson, Arizona, May 1997. | | | | |
| | F. Naumann and C. Rolker. <u>Do Metadata Models Meet IQ Requirements?</u> In <i>Proceedings of the International Conference on Data Quality (IQ)</i> , MIT, Cambridge, 1999. | | | | |
| | E. Rahm and H. Hai Do. <u>Data Cleaning: Problems and Current Approaches</u> . <i>IEEE Data Engineering Bulletin</i> , 23(4):3-13, December 2000. | | | | |
| | V. Raman and J. Hellerstein. <u>Potter's Wheel: An Interactive Data Cleaning System</u> . <i>VLDB 2001</i> , pages 381-390, Roma, Italy. | | | | |
| | S. Sarawagi and A. Ghani. <u>Interactive Deduplication Using Active Learning</u> . In <i>Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery in Databases</i> , Edmonton, Canada, July 23-26, 2002. | | | | |
| EXAMINER:  | | | | | |
| DATE CONSIDERED 12/21/05 | | | | | |
| EXAMINER: Initial if citation considered, whether or not citation is in conformance with MPEP § 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant. | | | | | |